

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Using Epistemic Synchronization Index (ESI) to measure students' knowledge elaboration process in CSCL

Ning Ding^{a,*}, Jieqiang Wei^b, Marca Wolfensberger^c^a International Business School, Hanze University of Applied Sciences, Groningen, The Netherlands^b Faculty of Mathematics, University of Groningen, The Netherlands^c Hanze University of Applied Sciences, Groningen, The Netherlands

ARTICLE INFO

Article history:

Received 11 April 2014

Received in revised form

4 August 2014

Accepted 7 August 2014

Available online 27 August 2014

Keywords:

CSCL

Epistemic engagement

Knowledge elaboration

Sequential analysis

Content analysis

ABSTRACT

Researchers in CSCL have used a wide range of qualitative and quantitative methods to track students' cognitive involvement during collaboration. However, neither individual method suffices the need to capture the dynamic evolvement of students' epistemic engagement in CSCL. We developed Epistemic Synchronization Index (ESI) to quantify students' epistemic engagement and evolvement. ESI reveals the knowledge elaboration process of groups, and it helps researchers as well as teachers to distinguish epistemic involvement between members within one group.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The potential of Computer-Supported Collaborative Learning (CSCL) in terms of information retaining and higher-order thinking skills has been clearly addressed in the previous research (e.g. [Garrison & Arbaugh, 2007](#)). The pedagogical essence of collaborative learning is that during discourse students are engaged in building or elaborating knowledge through coordinated efforts ([Dillenbourg, Baker, Blaye & O'Malley, 1996](#)). Therefore, [Stahl and Hesse \(2006\)](#) claimed that the term “intersubjective meaning-making” seems to be a more appropriate description than “collaborative learning”. During the process of meaning-making, a group is a learning entity working towards a problem solution, and any individual work during collaboration is more or less interrelated with the other group members. Hence, simply looking into single group member's work is not sufficient to understand how the group jointly elaborates the knowledge and step by step arrives at the final solution ([Salomon, 1993](#)). The effectiveness of group problem-solving relies on individual contributions and the responding efforts from his/her group members.

Echoing the introduction of measuring students' collaborative problem-solving skills through computer-based interactions in the international PISA test in 2015, education researchers as well as policy makers have been paying an increasing attention to the theoretical and methodological challenges of knowledge assessment in CSCL ([Stahl, Law, & Hesse, 2013](#)). The current research is an exploratory study based on a theoretical notion of our previous studies ([Ding, 2009](#); [Ding, Bosker, & Harskamp, 2010](#)). It aims at exploring a method to measure the synchronization of students' epistemic involvement at CSCL. It is named as Epistemic Synchronization Index (ESI). In the following sections, we are going to introduce the theoretical background and development of ESI, followed by a description of the study methods. Subsequently, we will explain the way to calculate ESI and introduce the preliminary results. After that, conclusions will be drawn and the theoretical and practical implications of ESI will be discussed.

* Corresponding author. International Business School (IBS), Hanze University of Applied Sciences, P.O. Box 70030, 9704 AA Groningen, The Netherlands. Tel.: +31 (0)50 595 2346, +31 (0)61 426 2533 (mobile); fax: +31 (0)50 595 2370.

E-mail address: n.ding@pl.hanze.nl (N. Ding).

URL: <http://www.hanzeuniversity.eu>.

2. Theoretical framework

2.1. Measuring epistemic engagement in CSCL

Previous research has shown that higher-order thinking can be triggered via interaction with peer learners during collaborative problem-solving (Duffy & Cunningham, 1996; Johnson & Johnson, 1993; Paus, Werner, & Jucks, 2012). In order to solve the problem, participants need to recall and process their knowledge (Hinsz, Tindale, & Vollrath, 1997), explicate their own ideas, and reflect on peer learners' contribution (Pena-Shaff & Nicholls, 2004). Muukkonen and Lakkala (2009) found that a high level epistemic engagement during problem-solving is essential for a group to achieve knowledge-creation. Larreamendy-Joerns and Leinhardt (2006) stated that online learning is a process of participatory practice within which students learn through social interaction and negotiation of meaning. In line with them, Shea and Bidjerano (2009) highlighted that learners are collaborative knowledge builders, and their ongoing understandings can be tracked by looking into their contributions to the conversation. And the conversation as a whole supports learning. Such kind of epistemic engagement includes recognizing the discrepancies of knowledge, reorganizing concepts and reconstructing conceptual understandings manifested during the discourse of collaborative problem-solving (King, 2007). From the perspective of epistemic engagement approach, in the current study, the conceptualization of epistemic engagement is based on the definition of epistemology of Sandoval (2005). They distinguished the epistemic understanding and epistemic practice. The former one is the knowledge that students acquire and apply in thinking. The latter one refers to the cognitive and discursive activities that students are involved in to develop their epistemic understanding. Weinberger and Fischer (2006) proposed to measure the amount of epistemic engagement such as task-related messages during online communication.

In the experimental research of Paus et al. (2012), psychology students at two German universities were randomly paired. Researchers focused on students' conceptual elaboration during dyadic online discourse. A positive relationship regarding conceptual elaboration was found between dyadic and individual levels. Their study suggested that conceptual explanations function as the most important indicator of a successful group problem-solving (Eryilmaz, van der Pol, Ryan, Clark, & Mary, 2013). While measuring the effectiveness of collaborative learning, these epistemic activities should be taken into account as important indicators.

There are many ways to measure students' epistemic engagements during collaborative problem-solving, and CSCL provides a rich corpus of data for researchers to gain insights into the ways in which students co-construct and elaborate their knowledge over time (Palmer, Holt, & Bray, 2008). Some researchers examined the length or number of on-task contributions and detected a positive relationship between these indicators and individual knowledge acquisition (e.g. Barab & Duffy, 2000; Weinberger & Fischer, 2006). More specifically, the more engagement of students during collaboration, the better knowledge elaboration they might experience.

Nowadays, much more attention has been directed to the *qualitative content analysis* of students' online artifacts (Zhao & Chan, 2014). Hmelo-Silver (2003) stressed the importance of fine-grained content analysis. In the previous research, content analysis has been defined as a generic method that includes comparing, contrasting, and categorizing textual data (Hara, Bonk, & Angeli, 2000). Qualitative analyses of students' online interactions may help researchers delve into students' cognitive thinking and knowledge elaboration process because students' online messages can be viewed as conceptual artifacts, and can be continually elaborated via progressive discourse with peer learners.

Researchers have been working on analyzing the quality of students' activities in various CSCL environments, such as their posts at the asynchronous online discussion forum (e.g. Zhao & Chan, 2014), behaviors during synchronous video-conferencing (e.g. Meier, Spada, & Rummel, 2007), or text-based messages in synchronous CSCL settings. Accordingly, various coding schemes have been explored from different perspectives. For instance, Kumpulainen and Mutanen (1999) distinguished three cognitive processing modes: *procedural processing*, *interpretative or exploratory processing* and an *off-task activity* mode. If students' cognitive work stayed at a surface level, it was regarded as a procedural processing. In contrast, if students elaborated their knowledge and contributed to the final solution of the problem in a meaningful manner, it was defined as an interpretative or exploratory processing.

From the conceptual and functional perspective, Fischer, Bruhn, Grasel, and Mandl (2002) established four collaborative knowledge construction processes: externalization, elicitation, integration-oriented, and social conflict-oriented consensus building. In this case, social interaction was taken into consideration. Muukkonen, Lakkala, and Hakkarainen (2005) proposed a coding scheme categorizing the main elements in the progressive inquiry model, that is, *Problem*, *Own explanation*, *Source-based explanation*, *Metacomment*, *Quote* or peer learner's idea, and *Reference to lecture*. From empirical study, the researchers have noticed that the category "own explanation" tended to simplify students' ideas. Therefore, Muukkonen and Lakkala (2009) refined the model and differentiated five elements of knowledge-creating inquiry practices, namely, *Question*, *Explicative knowledge*, *Theoretical knowledge*, *Metaknowledge*, and *Organize*. Focusing on three-student groups, they found that a high level of epistemic involvement in knowledge advancement was essential for knowledge-creating practice. They proposed three encompassing aspects of inquiry process: individual efforts, collaboration in group, and the development of the high-quality knowledge objects.

Based on the qualitative content analysis of students' communication artifacts during collaboration, researchers are also making an effort to *quantify the categorizations* in order to capture an overview of students' epistemic engagements. For instance, Hakkarainen (2003) focused on the epistemological nature of knowledge and differentiated five levels in the study with elementary school students in biology classes about progressive inquiry: Level 1: Separated pieces of facts; Level 2: Partially organized facts; Level 3: Well-organized facts; Level 4: Partial explanation; Level 5: Explanation. For each level a numerical score was assigned, ranging from 1 to 5.

Meier et al. (2007) have quantified the quality of collaborative process instead of counting frequency. They established a rating scheme consisting of five aspects of the collaborative process and the resulting nine dimensions, Communication (sustaining mutual understanding, dialogue management), Joint information process (information pooling, reaching consensus), Coordination (task division, time management, technical coordination), Interpersonal relationship (reciprocal interaction), and Motivation (individual task orientation). They looked into 40 dyads' performance during videotaped collaboration. The dimension of "individual task orientation" ranged from −2 (very bad) to +2 (very good).

Such kind of granularity gleaned from the enormous corpus of data depicts a meaningful picture for us to understand how knowledge is created or elaborated in a group. Dillenbourg (1999) defined three types of symmetry dimensions of collaborative learning: *the symmetry of action*, *the symmetry of knowledge*, and *the symmetry of status*. In light of the focus of current research on the process of knowledge

elaboration during collaborative problem-solving, we concentrated on the *symmetry of knowledge*, which refers to the level of knowledge and expertise among participants during collaborative work. Sangin, Molinari, Nuessli and Dillenbourg (2008) conducted an experimental study on 32 dyads in a synchronous collaboration setting with the aim to explore the effectiveness of a visualization tool of knowledge awareness (KAT). An interaction effect was found between the application of KAT and the pair-asymmetry factors concerning knowledge. In their study, symmetry of knowledge was measured through pre- and post-test, which means, at the outset and the end of the collaboration. However, as Reimann (2009) pointed out, the crucial element in CSCL is the process because “learning unfolds over time”. The current theoretical constructs and methods in CSCL research tend to neglect the information of time and sequence. Against this background, *sequential analysis* of students' communication artifacts has gained much attention in CSCL research.

Based on three cognitive modes proposed by Kumpulainen and Mutanen (1999), Ding (2009) and Ding et al. (2010) developed a coding system to measure students' problem-solving process, which was termed “elaboration values” referring *off-task*, *on-task* and *elaboration* activity. In synchronous CSCL, each piece of student online messages was coded into a discrete numerical value as -1 , 0 , or $+1$. As Muukkonen and Lakkala (2009) claimed, CSCL researchers should pay attention to the epistemic involvement of individual participant and the group as a learning community as well. They viewed the collaborative problem-solving as a cumulative process on the basis of individual contributions. We cumulated students' elaboration values and arrived as three visualized patterns of knowledge elaboration illustrating the process of dyadic knowledge elaboration process. The *divergent* pattern featured by two diverging curves shows an increasing cognitive discrepancy between two participants. The *cross* pattern illustrates that students' knowledge elaboration processes are closely intertwined. The participants keep a close eye on their partner's processing and take turns in dominating the knowledge elaboration. The *parallel* pattern shows two roughly parallel curves, indicating that the cognitive gap between the two participants stays the same during collaboration. The patterns present a general picture of how students within one dyad elaborate on knowledge sequentially.

With deeper delving into CSCL data in the past three years, we've recognized the limits of these visualized patterns. For instance, the determination of patterns was not normalized, which indicates that categorizing the process into patterns heavily relied on individual judgment. The distinction among the patterns was ambiguous and the categorization could be subjective. Moreover, we argued that the time sequence of the artifacts needs to be taken into consideration while measuring the effectiveness at the group level. As Furberg, Kluge, and Ludvigsen (2013) stressed, students' meaning-making process is an evolving process consisting of moment-to-moment interactions. Tracking students' activities over time may help us to depict how the problem-solving process unfolds via joint efforts (Engle & Conant, 2002).

To date, very little research on individual epistemic involvement suffices the need to capture the dynamic progress of the evolvement of individual epistemic engagement in a group setting. It is difficult for us researchers to grasp the overall process of students' epistemic engagement in CSCL by simply tallying the frequency of messages. Against this background, we advanced our previous method using elaboration values and explored a new method, calculating *Epistemic Synchronization Index (ESI)*, to measure the degree of synchronization of students' epistemic involvement during collaborative problem-solving.

2.2. Epistemic Synchronization Index (ESI)

Derived from Dillenbourg's notion of the symmetry of knowledge, Epistemic Synchronization Index (ESI) was developed to measure whether the participants are on the same page of epistemic engagement during collaborative work. We concentrated on students' synchronous and text-based collaborative problem-solving, and proposed to take the time sequence into consideration.

The rationale for this proposition is based on the following assumptions:

1. Individual knowledge elaboration process needs to be highlighted and embedded against the background of group problem-solving process. Schegloff (2007) stressed that understanding the positioning of utterances or actions and how they were organized is essential.
2. Students' online text-based messages can be taken as artifacts on the belief that these messages can reflect their epistemic involvement during collaborative problem-solving. As Kapur (2011) noted, the epistemic engagement of individual student is invoked by explanations or critiques emerging during group discourse. This notion suggested that group member's epistemic engagements are not interdependent, but they are interrelated and co-evolve over time.
3. Time sequence needs to be taken into account when we measure the effectiveness of collaborative problem-solving. According to Reimann (2009), “human learning is inherently cumulative”. The sequence of students' online messages for communication does not only reflect their individual thinking flow, but also indicate how these artifacts influence each other's epistemic engagement, presenting the learning process of a group.

The current study is an exploratory study, aiming at exploring a method to measure the dynamic group knowledge elaboration process.

3. Research context

3.1. Subjects and CSCL setting

The study was conducted in an International Business and Management School of a university in The Netherlands. Two female bachelor students from the fourth year participated in seven online collaboration sessions. One was from Ukraine and the other one was from China. Both of them have been following the English-taught courses of international business in The Netherlands for three years. Using English to communicate was no longer a challenge for both students. This dyad was chosen based on the following considerations. First of all, we focused on a female–female gender, trying to rule out the gender noise which may influence students' communication in CSCL. In our previous research, we've noticed that the communication style varies across gender during the process of problem solving (Ding, 2009). Secondly, for both students, English was not their native language. In the study by Lee (2004), native-speaking partners had certain difficulty and disaffection while chatting with a non-native speaker. We intentionally chose two students who spoke English as the second language. Thirdly, both students had a strong willingness to brush up their statistics knowledge because they are suffering pressure from the

individual project in business research. Finally, since they were selected from the same class, they knew each other very well before their collaboration. In the face-to-face classroom settings, these two students have had collaborated in several group projects in the past two years. Thus, they were quite familiar with each other's communication and working styles.

We used the course of Advanced Business Statistics for the second year bachelor students as an example. Both students had followed the course when they were in the second study year, and they passed the exam with sufficient grades, 7.5. However, they were busy with their internship in the third year, haven't been engaged into statistics any more. The study served as a refreshing course of their knowledge.

The study was carried out in an online chatting room provided by the school Blackboard system, which is a java-based and text-only chat application. Researchers used the instructor function to record all of students' text-based messages as well as the time slots for each message. The voice function of the chatting room was disabled in order to urge students to communicate via text-based messages.

3.2. Implementation of the study

Prior to the study, students received a brief training about how to enter the synchronous communication platform, the Chatting Room in the school Blackboard system, and how to type statistical symbols while online chatting. During the study, they were separated to different locations to avoid face-to-face communication. For each session, they were asked to watch a 20-min long video clip to brush up knowledge of statistical concepts and theories. Then, they received one question as well as hints via email. Doing so was to guarantee that both students started reading the problem at the same time. Once they received the questions, they could log into the online chatting room to discuss how to solve the problem. Once they agreed on an answer, they reported to the researcher in the chatting room. If the answer was correct, they could move to the next question. In case the answer was wrong, they could re-do it until it was correct. We assigned 2 h for each session. Students were allowed to finish problem solving within 2 h and left the session earlier. When they could not solve the problem within 2 h, they were required to stay in the chatting room until they arrived at the correct answer. Table 1 was the question used in the second CSCL session.

JetFly Aircraft Co. is developing a new wing de-icer system, which it has installed on 30 commercial airliners. The system is designed so that the percentage of ice removed is normally distributed with mean 96 and standard deviation 7.
The FAA will do spot check of six of the airplanes with the new system, and will approve the system if at least 98 percent of the ice is removed on average. What is the probability that the system receives FAA approval.

In our previous study regarding collaborative problem-solving of physics questions in a face-to-face setting, we've noticed that if interlocutors received the same hints and the hints were very well-structured and detailed, they might feel unnecessary to communicate with each other. Some students tended to solve the problem independently with the help of the hints. Therefore, we designed the hints with different content and most hints were question-based. Instead of instructing students what and how to do, we just used the hints to point at the direction of solution. Below (Table 2) are the hints we provided to students for the aforementioned question of Jetfly Aircraft.

3.3. Data analysis

The unit of analysis was defined as each message emerging at a recorded time slot. Before setting out to scrutinize the content of messages, we distinguished three epistemic levels of knowledge elaboration: off-task, on-task and elaborative messages (Ding, 2009; Ding et al., 2010). These three levels originated from the proposition of Kumpulainen and Mutanen (1999). In Table 3, we presented the coding categories as well as the examples. All messages emerging during seven sessions were coded into numerical values, either −1, 0 or +1. These were named as *Elaboration Value (EV)*.

For all on-task messages which representing students' epistemic engagement during collaboration, we distinguished *simple on-task* and *elaboration* messages. The former one referred to messages merely requesting, repeating or paraphrasing problem information, while the latter one indicated that students were trying to analyze the problem, integrate prior knowledge into new problem setting, and construct new knowledge. According to Hogan and Tudge (1999), to achieve a high level learning in a group context, the discourse within the group must be of a comparably high cognitive level. Such discourse would consist of, for example, thought-provoking questions, explanations, speculations, justifications, inferences, hypotheses, and conclusions.

Another essential aspect concerning coding needs to be noted. In text-based CSCL, the online interactions have some unique linguistic characteristics. For instance, the meaning delivered by students may be very fragmented. In face-to-face communication, speaker's messages are shaped by the previous turn of talk, showing that he/she understands the preceding turn. In contrast, there are loads of fragmented utterances over time in CSCL (Reimann, 2009). From our previous research (Ding, 2009; Ding et al., 2010), we have noticed that coding

Table 1
Sample statistics problem in the 2nd session of the study.

Worked-out example:	
Step 1:	$\mu = 96; \sigma = 7; n = 6$
Step 2:	$\therefore 6/30 = 0.2 \rightarrow 0.2 < 0.5$ \therefore Not necessary to use Finite Population Multiplier.
Step 3:	$\sigma_{\bar{x}} = \sigma / \sqrt{n} = 7 / \sqrt{6} = 2.835$
Step 4:	$A = \frac{1}{\int_{-0.5}^{0.5} ((CEV_{d1} + CEV_{d2})) * dt}$ $= \Rightarrow P = 0.2794$ $0.5 - 0.2794 = 0.2206$ $\therefore 22.06\%$ that the system will receive FAA approval.
Step 5:	

Table 2
Sample of hints.

	Student A	Student B
Hint 1	Do you know the sample size?	What is the population size?
Hint 2	Is the population in this case a finite population?	How to determine whether it is a finite or infinite population?
Hint 3	Visualization of the data is useful for your understanding.	The first step should start with the calculation of standard error.
Hint 4	Did you get a number, 2.60?	Check the z-score and probability.
Hint 5	Please review your problem solving process and see whether it makes sense.	How to interpret your final answer?

Table 3
Sample of the coding of epistemic values.

Category	Numerical value	Definition	Example
Off-task	−1	Content of the message is not related to the question solving.	<i>What is your plan for the Christmas vacation?</i>
On-task	0	<ul style="list-style-type: none"> - The message only shows an agreement, however, without justification or critical thinking. - The message is relevant to the question, however, without any advancement of solution. 	<ul style="list-style-type: none"> - <i>If we visualize it, each part should be 0.4722.</i> - <i>You mean 47.22%?</i> - <i>Ok, we know the z score now. Then, what shall we do?</i>
Elaboration	+1	<ul style="list-style-type: none"> - The message shows a step towards the final solution. 	<ul style="list-style-type: none"> - <i>It is a finite population because n/N is 0.2, smaller than 0.5.</i>

should embed each single message into the context, connecting it with its preceding and following turns. Table 4 is an example of our coding work.

Within 1 min and 23 s, this dyad has produced nine messages. Student 1 (S1) alone generated eight messages. She tended to use quite short and brief messages to deliver her ideas. Taking these fragmented pieces of meaning-making into account, we may find that simply counting the frequency of the messages could hardly inform us how the epistemic engagement of S1 evolved. Using our previous coding system (Ding, 2009; Ding et al., 2010), we avoided isolating the messages from their context. For example, message No.5 and 6 delivered the same meaning. No.6 presented the correct answer. But it was just a self-correction of the decimal, and should be deemed as one piece of information with No.5.

4. Results

During the seven sessions, this dyad has received seven statistics questions. We tallied the length of working time for each session and the amount of messages emerging in the chatting room. The descriptive data is shown in the Table 5.

To measure the synchronization of students' epistemic engagement over time, firstly, we coded each message into *Elaboration Value (EV)*, which was −1, 0 or +1. Then, we calculated the *Cumulative Elaboration Value* (hereafter “CEV”) for each individual student (s1 and s2) within the dyad, and plotted the CEV alongside the time sequence as Fig. 1 shows. The middle axis served as a baseline to split two students' CEV.

Subsequently, we segmented the total process of problem-solving equally into several time intervals. For each time interval, we calculated the area between individual CEV curve and the baseline. After that, we used the following formula to achieve the ESI value.

$$ESI_0 = \int_{t=0}^t A * |CEV_{s1} - CEV_{s2}| * dt \quad (1)$$

Where A is a normalizer defined as

$$A = \frac{1}{\int_{t=0}^t (|CEV_{s1} + CEV_{s2}|) * dt} \quad (2)$$

Table 4
Example of coding students' online interactions.

Message number	Time slot	Student	Message	Coding	Reasoning
1	0:01:12	1	X is between 27 and 29	0	On-task, repetition of problem information
2	0:01:35	1	I think	0	On-task
3	0:02:10	1	because it already gave us the P,	+1	Justification of the argument
4	0:02:43	1	it is 95.44%	+1	Extension of the previous justification
5	0:02:56	1	So if we visualize, each part should be 0/4722	+1	Proposal of solution
6	0:03:14	1	0.4722	0	Correction of preceding message
7	0:03:37	2	yes and z = 2	+1	Agreement and proposal to solution
8	0:03:49	1	yes	0	Agreement
9	0:04:07	1	and then we can find the standard error	+1	Proposal to solution

Table 5

Descriptive summary of students' online messages during seven CSCL sessions.

Session	Duration	Student 1	Student 2	Number of messages
1	1:03:55	95	83	178
2	0:35:00	63	54	117
3	1:38:43	134	125	259
4	0:40:21	58	53	101
5	0:49:26	68	56	124
6	0:36:24	60	58	118
7	0:29:09	47	42	89
Total	5:23:49	467	418	885

In order to calculate the A , the polygons of the CEV for individual students were defined as CEV_{s1} and CEV_{s2} . The integral function was applied in this case with the aim to achieve a more accurate comparison.

The closer ESI value approaches to 0, the more synchronized epistemic engagements can be expected between these two students. Suppose the ESI equals to 0, which is almost impossible for real collaboration setting, we may see two symmetrical CEV curves. In contrast, the larger value of ESI, the less synchronized epistemic engagement in collaboration.

Furthermore, we could use following two indexes, Equations (3) and (4), to measure whether an individual student contributes more to the collaboration than his/her partner.

$$ESI_1 = \int_{t=0}^t A * \max \{0, CEV_{s1} - CEV_{s2}\} * dt \quad (3)$$

$$ESI_2 = \int_{t=0}^t A * \max \{0, CEV_{s2} - CEV_{s1}\} * dt \quad (4)$$

Notice that we have $ESI_0 = ESI_{s1} + ESI_{s2}$. If $ESI_{s1} > ESI_{s2}$, the first student makes more contribution than the second student to the collaboration, and vice versa. The results were listed in Table 6. Having quantified all students' online messages emerging in seven sessions, we received the following findings of ESI for the dyads, which are shown in bold, and the ESI_{s1} , ESI_{s2} for individual students.

Now let's take Question 1 and 4 as examples to see what we may interpret from ESI. Q1 showed that both students started their epistemic involvement at a quite late stage, and the final ESI for this dyad was 0.5140, while the ESI of Q4 was higher, with 0.8378. Fig. 2 was the visualizations of ESI for Q1 and Q4, respectively.

When these two students were working on Q1, their epistemic involvement started at a quite late stage. Zooming in their online messages from 0:00:00 to 0:40:00, we found that these two students spent a lot of time on reading the problem and coordinating how to exchange information. One possible reason might be that this was their first time to work together in CSCL setting. In the following sessions, they set out right after they received the tasks.

Once they started, they elaborated their knowledge and solved the problem step by step, at a similar rate. The visualization of their epistemic involvement looked like a funnel. Such kind of epistemic synchronization was also manifested by the individual elaboration values. ESI_1 was 0.5405, very close to ESI_2 , 0.4595.

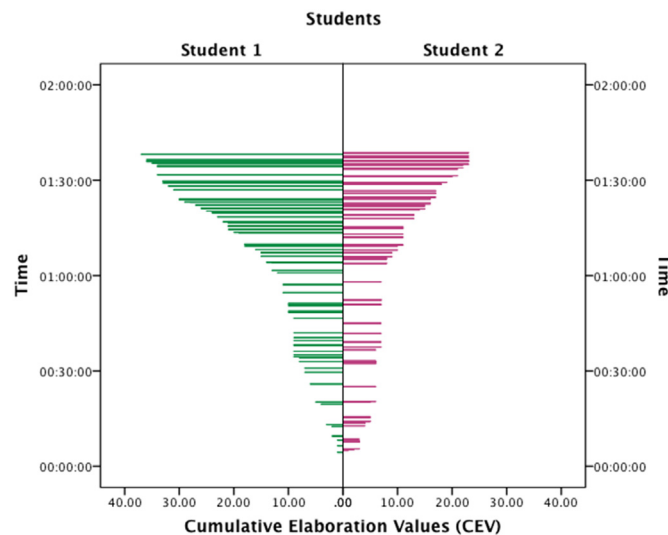
**Fig. 1.** Visualization of students' Cumulative Elaboration Values (CEV).

Table 6
Summary of students' *Epistemic Synchronization Index (ESI)* values*.

Question number	Duration	A	ESI	ESI ₁	ESI ₂
1	1:03:55	0.0270	0.5140	0.5405	0.4595
2	0:35:00	0.0238	0.7619	0.6667	0.3333
3	1:38:43	0.0163	0.7097	0.6613	0.3387
4	0:40:21	0.0270	0.8378	0.7838	0.2162
5	0:49:26	0.0385	0.7692	0.6154	0.3846
6	0:36:24	0.0263	0.6842	0.5789	0.4211
7	0:29:09	0.0303	0.7576	0.6667	0.3333

In Q4, students started their epistemic engagement at an early stage. From the visualization, it can be noticed that Student 1 tended to get involved more at epistemic level during collaboration, and the ESI₁ was 0.7838. In contrast, ESI₂ which was the ESI value of Student 2, was only 0.2162. In other words, the epistemic engagement of S1 was roughly three times more than that of S2, $0.7838/0.2162 = 3.63$. In this case, the funnel-like visualization presented an asymmetrical shape.

5. Conclusion and discussion

Collaborative learning is not bound to be productive, especially when computers mediate the communication. In CSCL, students' interaction can be deemed as artifacts reflecting students' state of knowledge and stance towards the task. In order to solve a problem, students' interaction with a high level of epistemic engagement is essential (Gorsky & Caspi, 2005; Muukkonen & Lakkala, 2009). Tracking students' epistemic activities over time may help us to depict how the problem-solving process unfolds via joint efforts (Engle & Conant, 2002).

Researchers in CSCL have used a wide range of qualitative and quantitative methods to track students' cognitive involvement during collaboration (Palmer et al., 2008). Quantitative methods included tallying the frequency of students' messages emerging in collaboration. Yet, as Palmer et al. (2008) pointed out that simply using quantitative analyses of students' interactions cannot yield a holistic picture of their cognitive and social engagement in the collaboration while qualitative analyses are open to variation of classification by different observers (see also Cotton & Yorke, 2006). In 2006, Strijbos, Martens, Prins, and Jochems published an article named "Content Analysis: What are they talking about?" in the journal *Computers & Education*. They called upon a critical reflection on quantitative content analysis of CSCL messages and triggering a discussion on methodological exploration. Previous qualitative analyses mainly focused on the content of students' online interactions in order to delve into students' cognitive thinking and knowledge elaboration process. However, Furberg et al. (2013) stressed that students' meaning-making process is an evolving process consisting of moment-to-moment interactions. Merely focusing on the content of messages may blur the sequence of students' thinking and the flow of ideas. So far neither individual method suffices the need to capture the dynamic evolvement of students' epistemic engagement in CSCL.

Aiming at grasping the dynamics of students' online collaboration regarding their epistemic engagement, we explored a new method, *Epistemic Synchronization Index (ESI)*, a combination of qualitative content analysis and quantitative measurement.

In order to capture a holistic picture of the evolvement of students' epistemic engagement during collaboration, firstly, we used a simple coding system to categorize students' online messages. Suthers and Xu (2002) defined students' online messages as "conversationally fragmented discourse contributions". A detailed categorization of messages runs the risk of crippling the integrity of information that messages delivered.

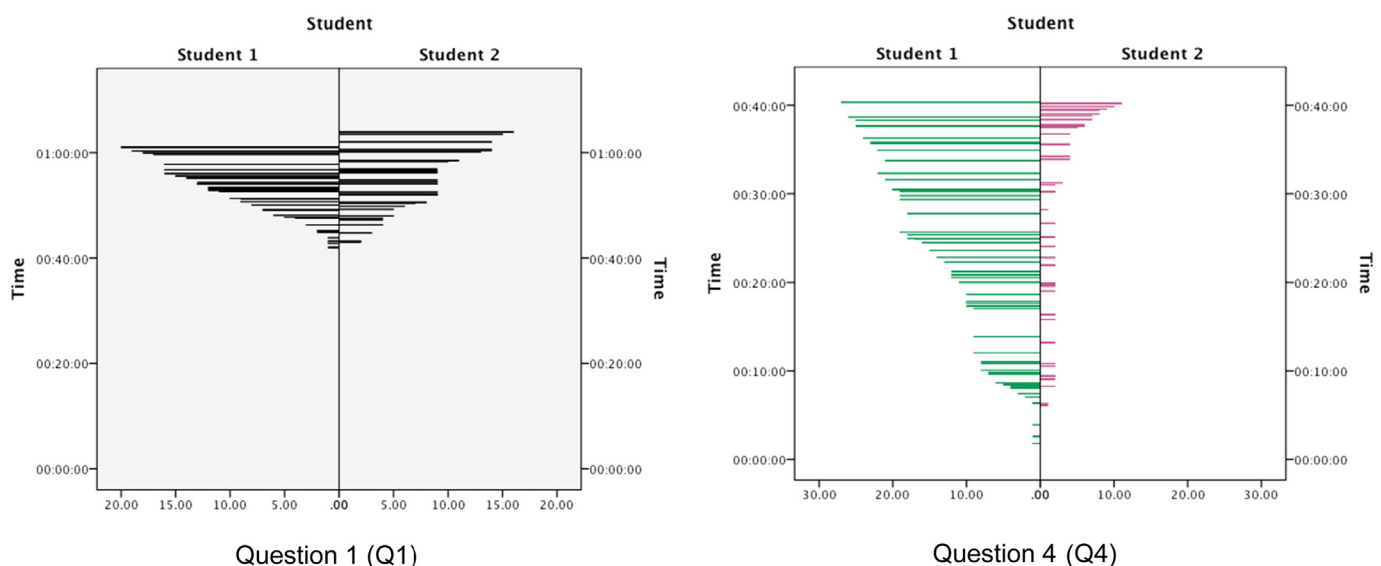


Fig. 2. Visualization of Epistemic Synchronization Index (ESI).

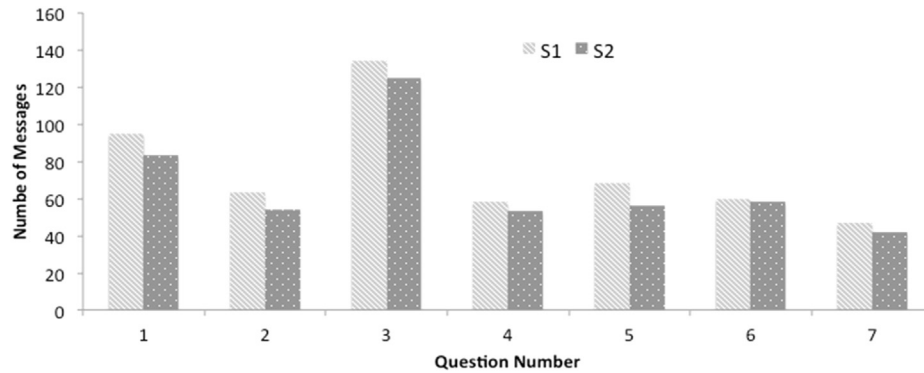


Fig. 3. Number of messages of student 1 (S1) and student 2 (S2).

Secondly, we summed up the values to measure students' epistemic involvement during collaboration for each individual student. Based on this cumulative values, we visualized the process of students' knowledge elaboration. This visualization has the potential to be presented in a real-time CSCL setting in the future. As Janssen, Erkens, Kanselaar, and Jaspers (2007) noted, using visualization of collaboration process may largely reduce the amount of off-task messages and increase the learning effectiveness. From our visualization of students' work, we can get a general picture whether the epistemic engagement during collaboration between two students is on the same page.

Furthermore, we developed ESI formulas. ESI is a value ranging from 0 to 1. The more unsynchronized collaboration at epistemic level, the higher the value of ESI. As we mentioned earlier, looking at the amount or frequency of online messages generated during CSCL or the final products could hardly reveal who contributed more and who might lag behind, or whether group members are working on the same page. In order to reinforce this argument, we visualized the summary statistics of Tables 5 and 6, respectively, as shown in Figs. 3 and 4. In Fig. 3, there was no obvious difference between the number of messages generated by S1 and S2. It indicated that, for this question, S1 and S2 produced a similar amount of online messages. Yet, looking at their individual ESI values in Fig. 4, we can easily detect that S1 engaged into a much deeper epistemic level than S2 throughout seven questions.

All in all, using ESI to measure students' epistemic engagement alongside the timeline in CSCL sheds light to the following three fields: (1) it quantifies students' epistemic engagement of CSCL and (2) it goes beyond the surface understanding of students' collaboration, and enables us to look into how knowledge develops over time and how it is co-built by collaborative students; and (3) instead of taking the group as a whole for analysis, it distinguishes the epistemic engagement of the individual student within a dyad.

In the future, we may generalize this method to answer various research questions. For example, we may apply it to conduct some comparative studies such as how the epistemic engagement during collaboration varies across gender, age groups, ethnicities etc. Moreover, we may explore the relationship between students' epistemic engagement and their academic performance using pre- and post-test design. Besides, we may zoom in the moment that students' epistemic engagement diverges to diagnose the factors that result in an unbalanced collaboration.

We believe that this methodological exploration will not only contribute to the CSCL theoretical research, but can also be applied into various online learning environments. For instance, using this algorithm we can design a real-time computer-based tutoring environment. Based on question content and difficulty level, teachers can prepare loads of hints as reservoir. Once the server computer detects an unbalanced epistemic involvement, student who has the lower ESI will receive a hint. In contrast, for the one who has taken the lead in collaboration regarding knowledge elaboration, it is not necessary to him/her to get a hint.

6. Limitation

As a methodological exploration, we used one dyad's online collaboration as a case for qualitative content analysis, and only seven questions were assigned. Limitations are evident. Hereby, we proposed an agenda for CSCL researchers in their future study. First of all, we

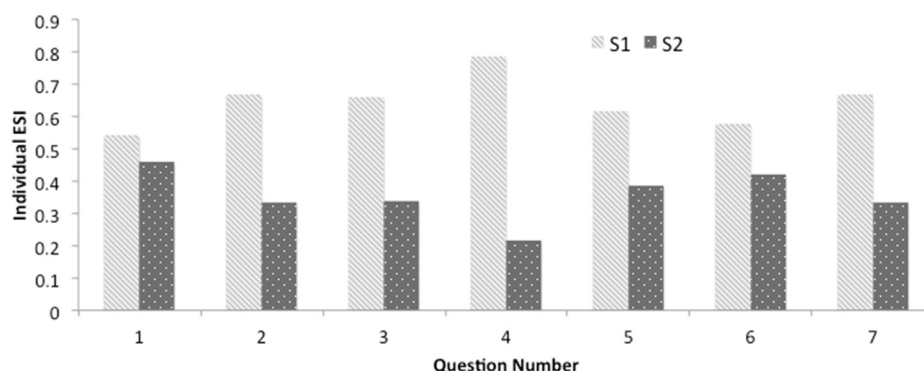


Fig. 4. ESI values for student 1 (ESI₁) and student 2 (ESI₂).

suggest to replicate the studies with much larger sample sizes in order to crystallize and validate the methods. In this regard, we suggest future researchers can take the quantitative approach into account. Then, we may explore the relationships between the Epistemic Synchronization Index (ESI) with the group learning achievement, and the individual ESI with students' individual learning gain through CSCL. In this regard, we recommend to use multilevel analysis to stratify students' performances at the group and individual level (Ding et al., 2010).

Moreover, categorizing the online messages needs a coordinated work among researchers. Currently, only the researchers analyzed students' online messages and categorized them. In the future research, more trained raters should be invited to analyze the content of online messages using the coding system. Their coding results of the online messages should be cross-checked. For a study with a large sample size, an inter-rater reliability should be presented along with the results.

Thirdly, in the current study, we focused on the business statistics which stresses students' interpretation and calculation skills. For other subjects such as physics or language learning, the categorization of students' epistemic engagement may vary to some extent. Therefore, researchers who want to replicate it in other subjects should be flexible in tailor the coding scheme to different subject content.

Last but not the least important, due to the task difficulty and students' knowledge level, we defined every minute as a time slot to calculate the ESI. The formulas leave a large room to adapt the time slots into different learning contexts. For instance, if there is a high frequency of exchanged messages between students, we can shorten the time slot and define every 30 s as a calculation unit.

References

- Barab, S. A., & Duffy, T. (2000). From practice fields to communities of practice. In D. Jonassen, & S. M. Land (Eds.), *Theoretical foundations of learning environments* (pp. 25–56). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Cotton, D., & Yorke, J. (2006). *Analysing online discussions: What are students learning?*. Paper presented at the 23rd Annual Asclite Conference: Who's learning? Whose technology? Sydney, Australia: Sydney University Press
- Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada, & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189–211). Oxford: Elsevier.
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–16). Switzerland: Elsevier.
- Ding, N. (2009). Visualizing the Sequential Process of Knowledge Elaboration in Computer-Supported Collaborative Problem Solving. *Computers & Education*, 52(2), 509–519.
- Ding, N., Bosker, R. J., & Harskamp, E. G. (2010). Exploring gender and gender pairing in the knowledge elaboration processes of students using computer-supported collaborative learning. *Computers & Education*, 56(2), 325–336.
- Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 170–198). New York: Simon & Shuster Macmillan.
- Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, 20(4), 399–483.
- Eryilmaz, E., van der Pol, J., Ryan, T., Clark, P. M., & Mary, J. (2013). Enhancing student knowledge acquisition from online learning conversations. *International Journal of Computer-Supported Collaborative Learning*, 8(1), 113–144.
- Fischer, F., Bruhn, J., Grasel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction*, 12, 213–232.
- Furberg, A., Kluge, A., & Ludvigsen, S. (2013). Student sensemaking with science diagrams in a computer-based setting. *International Journal of Computer-Supported Collaborative Learning*, 8(1), 48–64.
- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: review, issues, and future directions. *Internet and Higher Education*, 10(3), 157–172.
- Gorsky, P., & Caspi, A. (2005). Dialogue: a theoretical framework for distance education instructional systems. *British Journal of Educational Technology*, 36(2), 137–144.
- Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of Research in Science Teaching*, 40(10), 1072–1088.
- Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional science*, 28(2), 115–152.
- Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conception of groups as information processors. *Psychological Bulletin*, 121, 43–64.
- Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: multiple methods for integrated understanding. *Computers & Education*, 41, 397–420.
- Hogan, D., & Tudge, J. (1999). Implications of Vygotsky's theory for peer learning. In A. O'Donnell, & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 39–65). Mahwah, NJ: Lawrence Erlbaum Associates.
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: does it contribute to successful computer-supported collaborative learning? *Computers & Education*, 49(4), 1037–1065.
- Johnson, D. W., & Johnson, R. T. (1993). What we know about cooperative learning at the college level. *Cooperative Learning*, 3(3), 17–18.
- Kapur, M. (2011). Temporality matters: advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer-Supported Collaborative Learning*, 6(1), 39–56.
- King, A. (2007). Scripting collaborative learning processes: a cognitive perspective. In F. Fischer, I. Kollar, H. Mandl, & J. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational and educational perspectives* (pp. 13–37). New York, NY: Springer.
- Kumpulainen, K., & Mutanen, M. (1999). The situated dynamics of peer group interaction: an introduction to an analytic framework. *Learning & Instruction*, 9(5), 449–473.
- Larremendy-Joerns, J., & Leinhardt, G. (2006). Going the distance with online education. *Review of Educational Research*, 76(4), 567–605.
- Lee, L. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the U.S. *Language Learning & Technology*, 8(1), 83–100.
- Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *Computer-Supported Collaborative Learning*, 2, 63–86.
- Muukkonen, H., Lakkala, M., & Hakkarainen, K. (2005). Technology-mediation and tutoring: how do they shape progressive inquiry discourse? *Journal of the Learning Sciences*, 14(4), 527–565.
- Muukkonen, H., & Lakkala, M. (2009). Exploring metaskills of knowledge-creating inquiry in higher education. *International Journal of Computer-Supported Collaborative Learning*, 4(2), 187–211.
- Palmer, S., Holt, D., & Bray, S. (2008). Does the discussion help? The impact of a formally assessed online discussion on final student results. *British Journal of Educational Technology*, 39(5), 847–858.
- Paus, E., Werner, C. S., & Jucks, R. (2012). Learning through online peer discourse: structural equation modeling points to the role of discourse activities in individual understanding. *Computers & Education*, 58, 1127–1137.
- Pena-Shaff, J., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board (BBS) discussions. *Computers & Education*, 42, 243–265.
- Reimann, P. (2009). Time is precious: variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257.
- Salomon, G. (1993). No distribution without individual's cognition: a dynamic interactional view. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 111–138). Cambridge, UK: Cambridge University Press.
- Sandoval, W. A. (2005). Understanding students' practical epistemologies and the influence on learning through inquiry. *Science Education*, 89(4), 634–656.
- Sangin, M., Molinari, G., Nüssli, M. A., & Dillenbourg, P. (2008). Knowing what the peer knows: the differential effect of knowledge awareness on collaborative learning performance of asymmetric pairs. *Lecture Notes in Computer Science*, 5192, 384–394.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis* (Vol. 1). Cambridge: Cambridge University Press.
- Shea, P., & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster "epistemic engagement" and "cognitive presence" in online education. *Computers & Education*, 52, 543–553.
- Stahl, G., & Hesse, F. (2006). Focusing on participation in group meaning making. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 311–313.

- Stahl, G., Law, N., & Hesse, F. (2013). Collaborative learning at CSCL 2013. *International Journal of Computer-Supported Collaborative Learning*, 8(3), 267–269.
- Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: what are they talking about? *Computers & Education*, 46, 29–48.
- Suthers, D., & Xu, J. (2002). Kukakuka: an online environment for Artifact-Centered discourse. In *Education track of the Eleventh World Wide Web Conference (WWW 2002)* (pp. 472–480). Honolulu, May 7–11, 2002. Accessed at <http://www2002.org/CDROM/alternate/2521>.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46, 71–95.
- Zhao, K., & Chan, C. (2014). Fostering collective and individual learning through knowledge building. *International Journal of Computer-Supported Collaborative Learning*, 9(1), 63–95.